View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

CellPress
Neuron
Previews

When is the Time for a Change? Decomposing Dynamic Learning Rates

Adrian G. Fischer[1] and Markus Ullsperger[1,2,3,*]
[1]Otto von Guericke University Magdeburg, Institute of Psychology II, 39106 Magdeburg, Germany
[2]Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, 6525 Nijmegen, the Netherlands
[3]Center for Behavioral Brain Sciences, 39106 Magdeburg, Germany
*Correspondence: markus.ullsperger@ovgu.de
http://dx.doi.org/10.1016/j.neuron.2014.10.050

Humans flexibly weight incoming evidence when updating beliefs and adjusting behavior. In the current issue of *Neuron*, McGuire et al. (2014) show how distinct neuronal correlates of main factors underlying this weighting converge on a common mechanism driving belief updates.

An essential question when evaluating evidence for one's beliefs in a changing world is whether or not new information that diverges from expectation is indicative of an underlying change or just reflects expectable variations. Distinguishing change from noise is necessary as it may require adjustments of behavior in many contexts, for example when deciding whether a person with subfebrile body temperature should be quarantined for a suspected highly contagious infection with poor prognosis, when determining whether extreme air and water temperatures indicate global warming, or when evaluating whether unexpected empirical findings result from noisy measurement methods or require a revision of the tested hypothesis.

The degree to which the deviation of an event from expectancy influences the updating of beliefs about the environment can be considered a learning rate. By varying the impact of an observation in an environment in which sudden change may occur, it is possible to drastically reduce trial and error learning costs (Courville et al., 2006). However, it has been a long-standing debate as to which factors may drive learning to become faster or slower.

McGuire et al. (2014) designed a task to directly measure learning rates by dividing belief updates by the visual prediction error—that is, the difference between expected and experienced outcome. Notably, in contrast to many previous studies, belief updates were directly reflected in quantifiable behavioral adjustments. In this task, while fMRI was recorded, subjects had to infer the position of a helicopter hidden in a cloudy sky. The goal was to catch bags being dropped from the helicopter by placing a virtual bucket underneath it. Some noise (e.g., reflecting turbulences) was added to the bags' falling trajectory. Thus, each outcome in itself deviated more or less from the expected position of the helicopter, but some especially unexpected outcomes may have indicated that the helicopter has moved to a novel position. Such major shifts required the update of the participant's belief of the helicopter's position and an according shift of the bucket. Bags randomly contained gold (reward) or stones (no reward). Employing computational modeling, the authors derived two normative factors from a Bayesian model that they identified as having major influences on learning rates. Figure 1 summarizes the main findings schematically by depicting the model-based and incidental influences on the learning rate, their neuronal correlates, and their convergence.

The first normative factor is the probability of the occurrence of a change point (CP; an actual change in helicopter's position), which is positively dependent on the general likelihood of changes (prior or accumulated knowledge about the volatility of the environment) and increases when an outcome is particularly unpredicted or surprising under the current belief. In general, change may be certain, for example when encountering a new situation, or very low in a familiar, stable setting (O'Reilly 2013). Importantly, converging lines of evidence indicate that the relationship between updating and surprise is nonlinear: updating of a belief is associated with distinct phenomena both behaviorally and neuronally over and above the increased surprise associated with it (O'Reilly et al., 2013). In the current study, brain activity that could be uniquely attributed to this factor was mainly seen in occipital visual areas even when controlling for eye movements that are associated with especially unexpected outcomes. A prominent theory holds that within the neural code the degree of unexpectedness and confidence into a signal is embedded (Friston et al., 2006), which is supported by the induction of confidence into visually guided choices by microstimulation in visual areas (Fetsch et al., 2014). Relatedly, presentation of visual stimuli associated with high uncertainty has also been demonstrated to increase activity in visual areas (Vilares et al., 2012). This activity could thus reflect a visual surprise signal factoring into the decision to update beliefs.

The second normative factor is the uncertainty associated with the belief in the current model, which the authors term relative uncertainty (RU), as irreducible or expected uncertainty also has to be accounted for. RU is conceptually closely related to a stimulus' associability and is directly related to the uncertainty of the current estimate (Pearce and Hall 1980) or, inversely, the number of previous observations in a predictable environment. Therefore, RU is high following detected CPs associated with major belief updates and decreases when more observations accumulate. Expectably, high RU was also related to higher learning rates tying behavior more closely to the preceding outcome and disregarding historic ones.

CrossMark

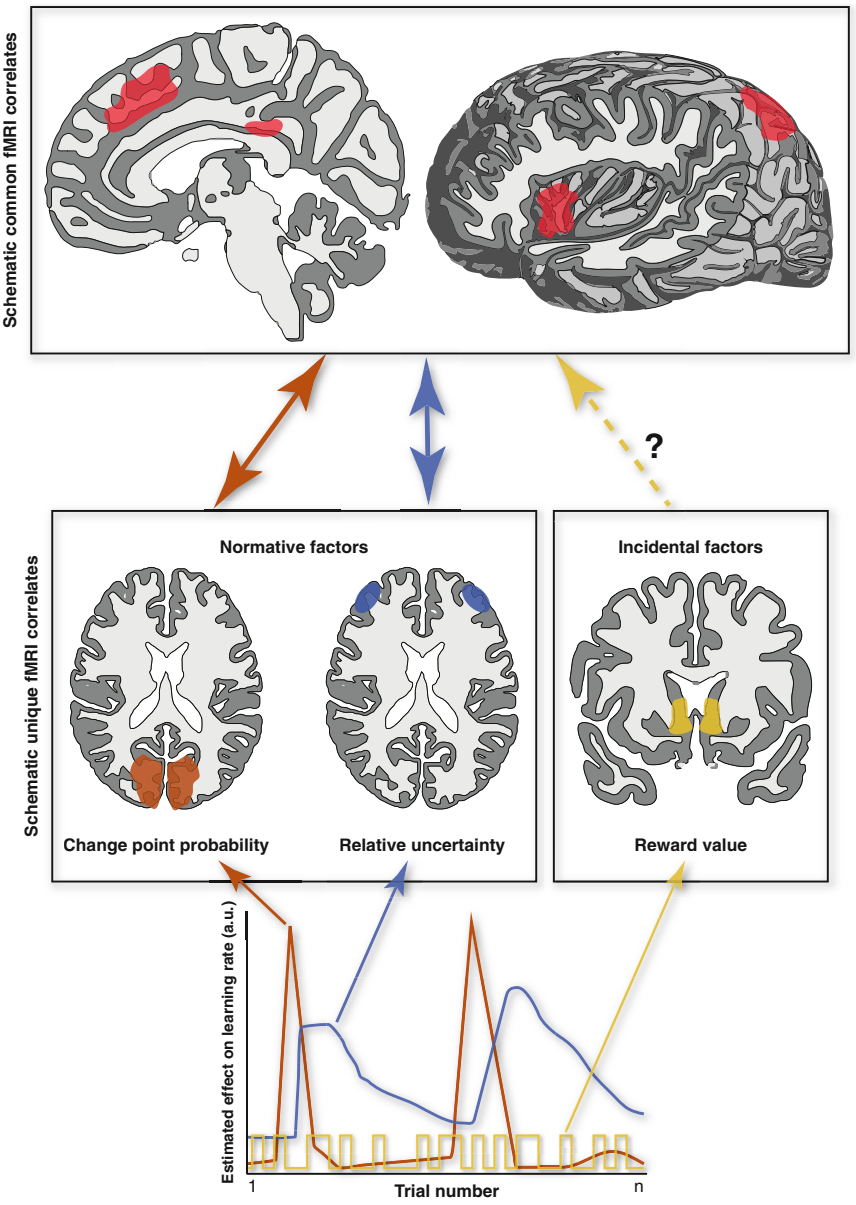

Figure 1. Relationship of Factors Influencing Learning Rates and Their Unique and Common fMRI Correlates

Bottom: CP probability (red) and RU (blue) both increase learning rates and are derived from a Bayesian model representing normative factors. CP probability increases quickly around CPs and rapidly decays, whereas RU shows a slower temporal dynamic. On rewarding trials, an additional learning rate increase is observed (yellow). Reward is randomly distributed across trials.

Middle: CP covaries uniquely with activity in bilateral visual cortices (left), whereas RU uniquely covaries with bilateral anterior PFC activity (middle). The reward regressor covaried with bilateral striatal BOLD responses.

Top: conjunct regions common to all three factors included DMFC, posterior cingulate cortices, parietal cortices, and bilateral anterior insula. Regions coding normative factors interacted with conjunct regions relative to their behavioral importance (double-headed arrows), while reward did not show such a clear effect (dashed arrow).

Analysis of fMRI correlates of this type of uncertainty showed unique increased activation in parietal cortex and bilateral anterior prefrontal cortex (aPFC). A very similar pattern was found in a study showing increased aPFC and parietal activity with increasing values of alternative unchosen options (Boorman et al., 2011). It will be a matter of future research to reconcile these results: are increased values of alternative options generally associated with higher RU, or does higher RU intensify the search and tracking of alternative options?

Other areas were less active when uncertainty was high. Decreased activity was found in ventromedial prefrontal cortex (vmPFC), which fits nicely to other studies that report value coding in vmPFC, as here the expected value of each trial following change points increases with increasing model precision.

Besides these normative factors, the analysis also revealed an influence of reward on the learning rate. Only gold bags were associated with reward, and their occurrence was randomized, but following these rewarding events learning rates also increased despite carrying no additional objective information. In the current task, this is a clear deviation from optimal behavior. As the authors point out, future work has to reveal whether a similar effect is found for punishing events. It should be noted, however, that in real life it is difficult to find rewarding or punishing events that are completely devoid of additional goal-related information, suggesting that a value effect on learning rate is in most cases adaptive. It could be speculated that this effect, reflected in bilateral striatal activity, is hardwired, i.e., Pavlovian.

Dynamic Influences on Common Learning-Rate Regions

However, McGuire et al. (2014) not only find activity unique for each factor, but also demonstrate conjunct activity that is influenced by all three behaviorally relevant factors. Most prominently (but not restricted to), activity in dorsomedial frontal cortex (DMFC, often also called posterior medial frontal cortex) was positively associated with all factors influencing learning rates. This fits well with observations where DMFC activity covaried with estimated volatility of the environment during outcome observation, and the degree of this covariation furthermore positively correlated with learning rates across subjects (Behrens et al., 2007). Furthermore, fMRI analysis of dynamic learning-rate magnitudes derived from a reinforcement learning model revealed a very similar region in DMFC (Krugel et al.,

2009). The findings also fit with the observation that lesions in a part of the DMFC, the anterior cingulate sulcus, result in deficits in modulating learning rate appropriately in a changing environment, thereby leading to severe impairments in reversal learning (Kennerley et al., 2006). The current study goes beyond the scope of previous studies by demonstrating very clear relationships between the activity in this common region and behavioral adaptation. The authors demonstrate that the coupling of regions specific for the probability of change points and RU with the common regions increases and decreases dynamically when each factor exerts more or less influence on learning rates behaviorally. Furthermore, a stronger reflection of the behavioral variables in the neural data also correlated with increased influences of the normative factors on behavior across subjects. In other words, the more similar the behavior of an individual was to the Bayesian model, the more these variables were reflected in the neural activity in the common learning-rate region. Although these fMRI data cannot elucidate the temporal dynamics of which region provides information first, these findings provide compelling evidence that a region within DMFC can adaptively shift the weight of different inputs from regions specialized in evaluating different factors to up- and downregulate the impact of external observations on beliefs about the state of a model. The pathway through which the value information is conveyed to the common region integrating influences on learning rate is still a matter of debate (Ullsperger et al., 2014), and future studies should address this interaction more directly.

Behavioral Residuals Explain Additional Neural Data

A very interesting aspect of the relationship between neural processes and behavioral adaptation lies within residual variance. Consider a situation in which a behavioral model predicts a participant's behavior with reliable accuracy. Still, on some trials, a subject may behave unexpectedly, for example by increasing the belief about the helicopter's position more than that predicted by the model. This additional variance can be used to identify neural correlates driving individual decisions that cannot be captured by a computational model (Fischer and Ullsperger 2013). McGuire et al. (2014) implemented the residuals of the behavioral analysis in the GLM of neural data and found that (a) to some degree higher activity in the common region was associated with additional behavioral effects and (b) unique activity in a DMFC region captured unexplained variance in behavior. This was further supported by a reciprocal analysis: single-trial BOLD estimates explained additional variance in the behavioral model. The first finding strengthens the link between behavior and fMRI activity, and the second finding hints to other factors that appear to have effects on learning rates that may not be captured by this behavioral model or in principle may be hard to account for. Although these findings have to be taken cum grano salis, the identified regions are compatible with salience coding regions, and activity here may reflect individual arousal fluctuating throughout the task.

In the current study, as in most other studies on the topic, change points were manifested as distinct jumps to a new position. However, for many natural scenarios, a gradual change may be more realistic. For example, when monitoring climate change, no manifest change point is likely to be found, but a change must be detected based on a consistent bias into a certain direction. Although some studies indicate that employing different models, that either assume a uniform distribution for new positions or drifts to the new state, appears to have comparable neural correlates (Behrens et al., 2007), this should be addressed explicitly. Furthermore, such a scenario would also enable the study of how subjects weight not only outcomes, but also changes in the underlying speed of drift that could accelerate or decelerate, providing even more naturalistic tasks.

In sum, McGuire et al. (2014) elegantly demonstrate how different variables represented in distinct brain regions are integrated by a common network to determine the right time to change ones beliefs by rapidly increasing learning rates and to adjust behavior accordingly.

REFERENCES

Behrens, T.E.J., Woolrich, M.W., Walton, M.E., and Rushworth, M.F. (2007). Nat. Neurosci. *10*, 1214–1221.

Boorman, E.D., Behrens, T.E., and Rushworth, M.F. (2011). PLoS Biol. *9*, e1001093.

Courville, A.C., Daw, N.D., and Touretzky, D.S. (2006). Trends Cogn. Sci. *10*, 294–300.

Fetsch, C.R., Kiani, R., Newsome, W.T., and Shadlen, M.N. (2014). Neuron *83*, 797–804.

Fischer, A.G., and Ullsperger, M. (2013). Neuron *79*, 1243–1255.

Friston, K., Kilner, J., and Harrison, L. (2006). J. Physiol. Paris *100*, 70–87.

Kennerley, S.W., Walton, M.E., Behrens, T.E., Buckley, M.J., and Rushworth, M.F. (2006). Nat. Neurosci. *9*, 940–947.

Krugel, L.K., Biele, G., Mohr, P.N., Li, S.C., and Heekeren, H.R. (2009). Proc. Natl. Acad. Sci. USA *106*, 17951–17956.

McGuire, J.T., Nassar, M.R., Gold, J.I., and Kable, J.W. (2014). Neuron *84*, this issue, 870–881.

O'Reilly, J.X. (2013). Front Neurosci *7*, 105.

O'Reilly, J.X., Schüffelgen, U., Cuell, S.F., Behrens, T.E., Mars, R.B., and Rushworth, M.F. (2013). Proc. Natl. Acad. Sci. USA *110*, E3660–E3669.

Pearce, J.M., and Hall, G. (1980). Psychol. Rev. *87*, 532–552.

Ullsperger, M., Fischer, A.G., Nigbur, R., and Endrass, T. (2014). Trends Cogn. Sci. *18*, 259–267.

Vilares, I., Howard, J.D., Fernandes, H.L., Gottfried, J.A., and Kording, K.P. (2012). Curr. Biol. *22*, 1641–1648.